

TextSpy 1.5

Příručka uživatele

Co je TextSpy	2
TextSpy Indexer.....	2
TextSpy Query	2
Použití TextSpy	
Vytváření indexu – TextSpy Indexer.....	3
Vyhledávání v indexech – TextSpy Query	5
Automatické připojení indexu TextSpy	9
Licenční politika pro TextSpy	10
Znamé problémy	10
Technická podpora	10
Další informace	11

Co je TextSpy

Amos Software TextSpy je řešení pro fulltextové vyhledávání pro Adobe Acrobat a jeho univerzální, multiplatformové dokumenty PDF. S jeho využitím lze indexovat kolekce souborů PDF a pak v nich velmi rychle vyhledávat požadovaná slova a slovní spojení. Toto řešení je postaveno na univerzální unicodové reprezentaci znaků a není tudíž omezené na práci s texty pouze v západní (Roman) či východoevropské znakové sadě.

Amos Software TextSpy, který vyvinula společnost Techsoft servis exkluzivně pro Amos Software, je sada zásuvných modulů (pluginů) pro Adobe Acrobat, která rozšiřuje jeho možnosti v práci s fulltextem o využití jazyků, které používají jiné znakové sady, než původní západní (Roman, Latin 1) jazyky. Hlavní rozdíl proti fulltextovému nástroji Search, je jazyková podpora, zahrnující kromě všech západoevropských jazyků i jazyky středoevropské a východoevropské, tedy jazyk český, slovenský, polský, slovinský, chorvatský, maďarský, rumunský, litevský, lotyšský, estonský, turecký, řecký, ruský, bulharský, ukrajinský a srbský.

Základem unicodového fulltextu TextSpy jsou dva moduly – Indexer a Query.

TextSpy Indexer

TextSpy Indexer vytváří ze zadaných dokumentů v univerzálním formátu Adobe PDF index s informacemi o výskytu a pozici jednotlivých slov. Ten je potom připraven pro použití s modulem TextSpy Query.

Poznámka: TextSpy Indexer se instaluje a pracuje pouze v Adobe Acrobatu 5.0 CE.

TextSpy Query

Jak už vyplývá z jeho názvu (query = dotaz), jde o dotazovací modul, instalovaný v Adobe Acrobatu 5.0 CE a Acrobat Readeru 5.0 CE, do něžž uživatel zadává dotazy na přítomnost daného slova nebo slov, a TextSpy Query potom pomocí indexu najde všechny výskyty, odpovídající zadaným kritériím. Tento výsledek je pak zobrazen jako seznam v okně, kde je možnost zobrazit informace o nalezeném souboru nebo zvolenou stránku otevřít v Adobe Acrobatu či Acrobat Readeru. Hledané výrazy jsou v otevřené stránce barevně zvýrazněny.

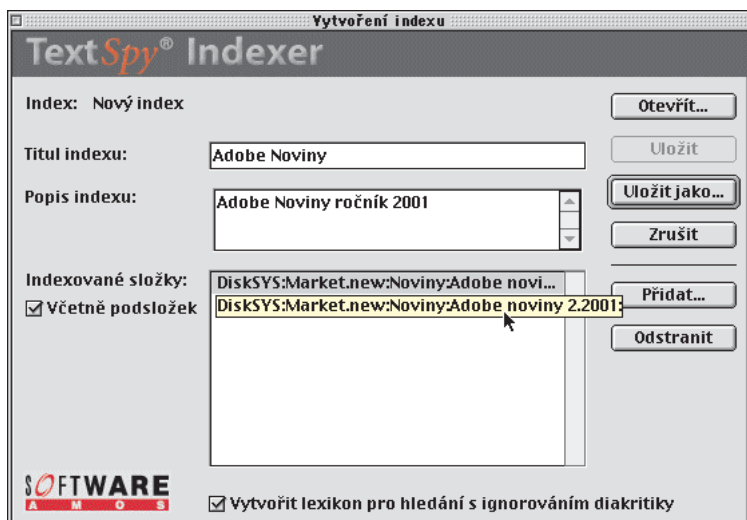
Na následujících stránkách najdete návod k použití TextSpy Indexeru a TextSpy Query.

Použití TextSpy

Vytváření indexu – TextSpy Indexer

Indexy se vytvářejí pomocí TextSpy Indexeru. TextSpy Indexer se vyvolá z běžící aplikace Adobe Acrobat 5.0 CE či vyšší z nabídky *Úpravy > TextSpy > Vytvořit index* (na Macintoshi též klávesovou zkratkou *Alt + M*).

Otevře se dialogové okno Indexeru.



Do dvou horních polí se zadává titlek a popis vytvářeného indexu, do spodního pole se přidávají (nebo se z něj odebírají) jednotlivé složky, obsahující soubory PDF, které mají být indexovány.

Pokud zaškrtnete volbu Vytvořit index prohledání s ignorováním diakritiky, Indexer navíc vytvoří soubory s daty pro zadávání dotazů bez diakritiky. To může být užitečné, pokud bude index využíván i na počítačích, kde není možné zadávat z klávesnice texty s diakritikou. Například, pro vyhledání slova „Škoda“ stačí zadat hledání výrazu „Skoda“.

Můžete vytvářet zcela nové indexy, nebo otevřít a přegenerovat či upravovat již existující indexy. Příkazem Uložit nebo Uložit jako... se spustí vlastní vytváření indexu (příkaz Uložit spustí aktualizaci již dříve vytvořeného indexu, příkaz Uložit jako... slouží k vytvoření a uložení nového indexu, nebo k aktualizaci již dříve vytvořeného indexu a jeho uložení pod jiným jménem nebo na jiné místo).

Volba „Včetně podsložek“ určuje, zda se mají indexovat pouze soubory umístěné přímo ve vybrané složce, nebo i soubory ve všech jejích podsložkách.

Při ukládání indexového souboru je třeba si uvědomit, zda je indexový soubor umístěn na stejném disku (svazku) jako indexované soubory, nebo na jiném. Zde je třeba v zásadě rozlišit dva různé způsoby využívání indexu:

1. Při používání TextSpy a jeho indexů pro vlastní potřebu, nebo v rámci pracovní skupiny v pevně dané počítačové síti je možné indexové soubory ukládat na libovolné místo (např. centrálně), bez ohledu na to, kde jsou soubory PDF, které chceme prohledávat. Pouze si musíme dát pozor, abychom později neměnili umístění indexovaných souborů PDF.
2. V případě vytváření indexů které mají být dále šířeny na různých přenosných médiích nebo přímo na CD-ROM, je nutné index a indexované soubory PDF umísťovat vždy v rámci jednoho svazku a při dalším šíření či kopírování zajistit, že jejich relativní pozice nebude změněna (je tedy možno přesouvat celé složky, obsahující index i indexované PDF soubory).

V podstatě jde o dodržení podobných pravidel, jako při tvorbě HTML stránek s odkazy.

Zejména v případě publikování na CD-ROM je z tohoto hlediska velmi vhodné zkontrolovat funkčnost vyhledávání dříve, než dojde k zadání výroby většího množství CD.

Pozor: Indexy vytvořené v této verzi TextSpy Indexeru není dovoleno šířit na veřejně distribuovaných discích CD-ROM nebo na jiných médiích. Více viz kapitola „Licenční politika pro TextSpy“.

Postup vytváření lze sledovat v následujícím okně.



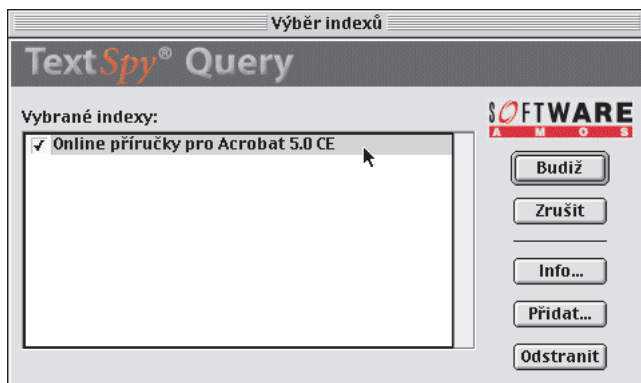
Horní indikátor ukazuje postup vytváření celého indexu, dolní zpracování jednotlivých souborů PDF. Hotový index je připraven k použití v TextSpy Query.

Vyhledávání v indexech – TextSpy Query



Vlastní vyhledávání se provádí v modulu TextSpy Query v Adobe Acrobatu 5.0 CE nebo v Acrobat Readeru 5.0 CE a novějších.

Nejprve je třeba určit indexy, ve kterých se má prohledávat. Ty se vyberou z nabídky *Úpravy > TextSpy > Vybrat indexy* (na Macintoshi též klávesová zkratka *Alt + X*).

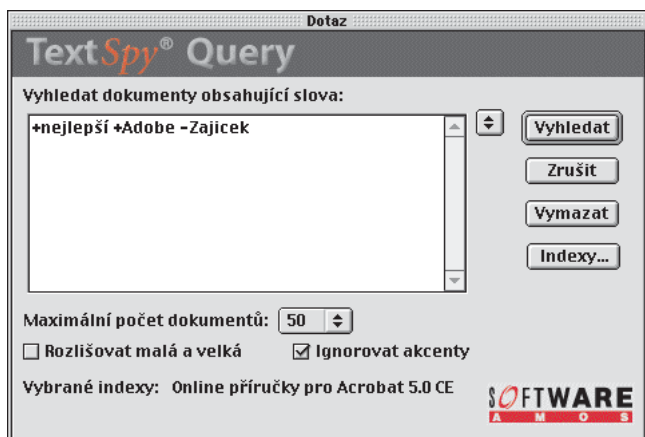
Otevře se dialogové okno.



V něm je možno přidávat, odebírat, aktivovat nebo deaktivovat indexové soubory, které mají být prohledávány. Po jejich vybrání je TextSpy Query připraven k vlastnímu vyhledávání. Pokud index není dostupný (např. byl-li připojen z CD-ROM), je řádek s jeho názvem šedivý, nedá se aktivovat a index se pro vyhledávání nepoužije.

Vlastní vyhledávání spustíte v Adobe Acrobatu 5.0 CE nebo Acrobat Readeru 5.0 CE či vyšším příkazem *Úpravy > TextSpy > Vyhledat*, klávesová zkratka *Ctrl + E* (Windows), resp. *Úpravy > TextSpy > Dotaz*, klávesová zkratka *Alt + F* (Macintosh) nebo z nástrojové lišty Acrobatu klepnutím na ikonu  (Windows), resp.  (Macintosh).

Otevře se dialogové okno.

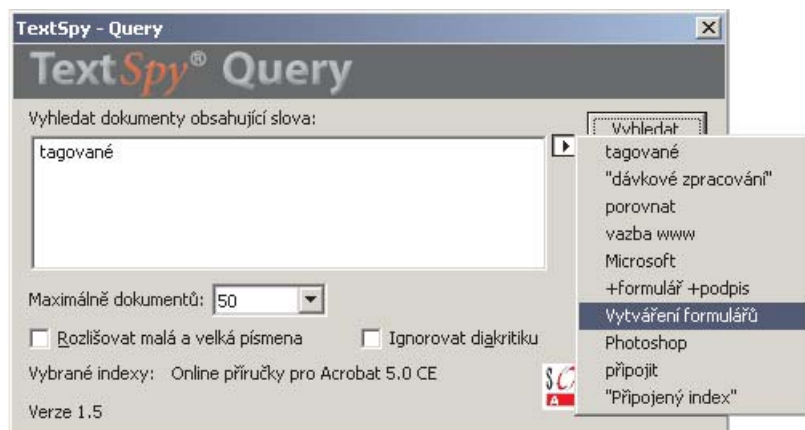


Vlastní vyhledávání je velmi intuitivní. Syntaxe dotazů není postavena přímo na booleovských dotazech, byl zvolen zápis obdobný jako v běžných internetových vyhledávacích. Do zadávacího okna se píše vyhledávaná slova (může jich být víc) a řídicí znaky pro následující případy:

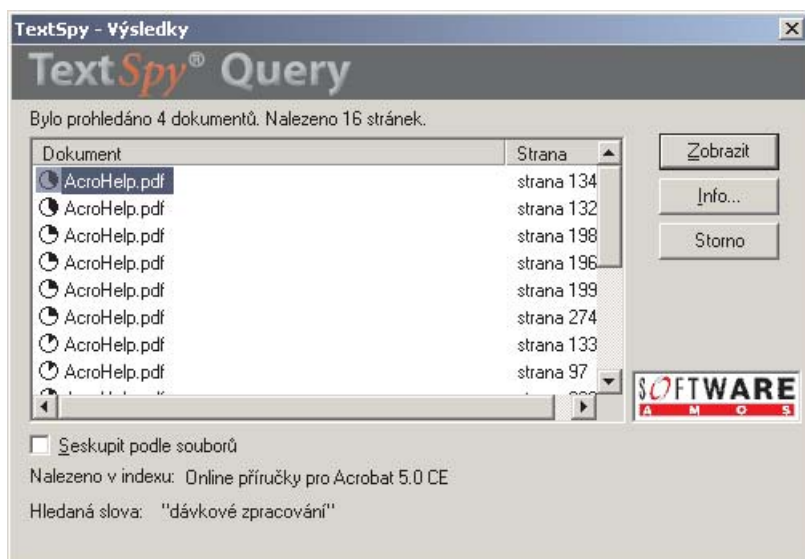
1. Vynucování nebo vylučování slov – výskyt slova ve výsledku dotazu lze vynutit umístěním znaku (+) přímo před slovo v okně dotazu. Aby výsledek nějaké slovo naopak neobsahoval, před dané slovo se umístí znaménko (-).
2. Použití zástupných znaků – zadáním hvězdičky (*) nebo otazníku (?) v zadávaném slově lze vyhledat více tvarů slova. Hvězdička zastupuje libovolný počet libovolných znaků, otazník jeden libovolný znak.
3. Rozlišování velkých a malých písmen – volí se zaškrtnutím příslušného okénka.
4. Ignorování diakritiky – volí se zaškrtnutím příslušného okénka. Tato volba se uplatní pouze v případě, že index byl vytvořen tak, aby obsahoval data pro vyhledávání bez diakritiky.
5. Vyhledávání přesných frází s více slovy – požadavek na vyhledání přesné fráze lze zadat jejím umístěním do uvozovek. Fráze se vyhledává jako posloupnost slov, bez ohledu na znaky, které jednotlivá slova oddělují.

Stručný popis syntaxe dotazů je rychle dostupný přímo z Acrobatu v okně *Nápověda* > *Co jsou zásuvné moduly* > *TextSpy Query...*

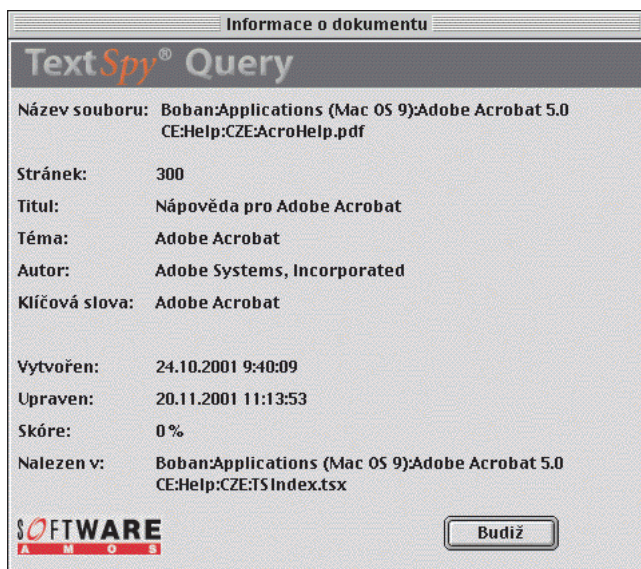
Několik (až 10) předchozích vyhledávání lze vyvolat pomocí nabídky na tlačítku vpravo od zadávacího okna:



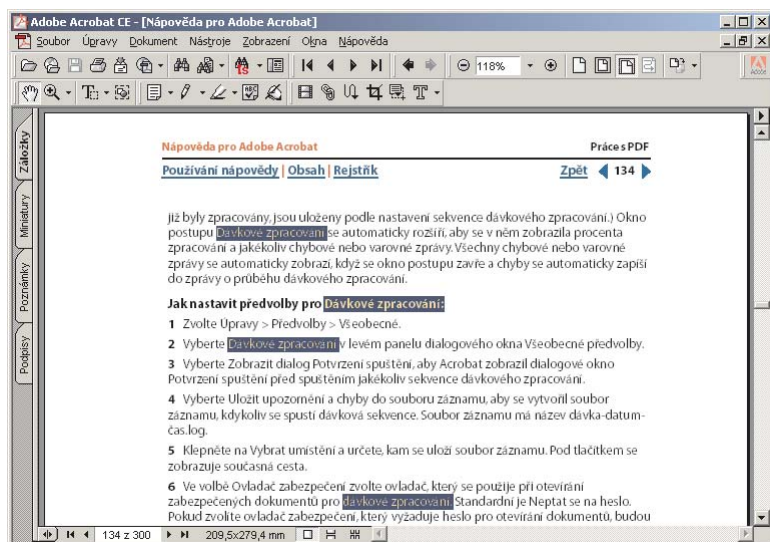
Po stisknutí tlačítka *Vyhledat* se otevře následující dialogové okno se seznamem nalezených stran:





V něm jsou seřazeny stránky, vyhovující zadaným kritériím podle váhy. V tomto základním okně výsledků je uveden název souboru a číslo stránky. Další informace lze vyvolat tak, že se zvolí příslušný řádek seznamu a klepne se na tlačítko *Info...*







Poklepáním na řádek v seznamu výsledků nebo klepnutím na tlačítko *Zobrazit* se daná stránka otevře v Adobe Acrobatu nebo Acrobat Readeru a zvýrazní se hledané výrazy.



K seznamu výsledků (nalezených stran) se lze vrátit buď příkazem *Úpravy > TextSpy > Výsledky*, popř. pomocí klávesové zkratky *Ctrl + ;* (Windows), resp. *Alt + G* (Macintosh),

nebo klepnutím na ikonu  (Windows), resp.  (Macintosh) v nástrojové liště Acrobatu.

V seznamu výsledků se lze rovněž pohybovat dopředu příkazem nabídky *Úpravy* > *TextSpy* > *Další stránka*, popř. pomocí klávesové zkratky *Ctrl* + > (Windows), resp. *Apple* + . (Macintosh), nebo klepnutím na ikonu  (Windows), resp.  (Macintosh) v nástrojové liště Acrobatu.

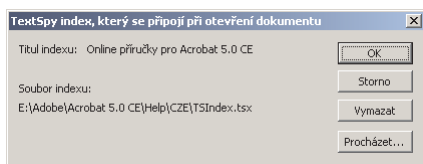
Pohyb zpět je analogický – příkazem *Úpravy* > *TextSpy* > *Předchozí stránka*, popř. pomocí klávesové zkratky *Ctrl* + < (Windows), resp. *Apple* + , (Macintosh), nebo klepnutím na ikonu  (Windows), resp.  (Macintosh) v nástrojové liště Acrobatu.

Možný je i pohyb mezi celými dokumenty (tedy nikoliv pouze po stránkách), ve směru dopředu příkazem *Úpravy* > *TextSpy* > *Další dokument*, popř. pomocí klávesové zkratky *Ctrl* + *Shift* + > (Windows), resp. *Shift* + *Apple* + . (Macintosh), ve směru zpět příkazem *Úpravy* > *TextSpy* > *Předchozí dokument*, popř. pomocí klávesové zkratky *Ctrl* + *Shift* + < (Windows), resp. *Shift* + *Apple* + , (Macintosh).

Automatické připojení indexu TextSpy

K PDF dokumentu je možné připojit informaci o souboru indexu, který se po otevření dokumentu automaticky přidá do seznamu prohledávaných indexů.

Chcete-li k PDF dokumentu připojit index, nebo chcete-li toto připojení změnit, zvolte v Adobe Acrobatu *Úpravy* > *TextSpy* > *Připojený index*. Objeví se dialogové okno s informacemi o připojeném indexu.



Klepněte na tlačítko *Procházet* a vyberte soubor indexu. Pak klepněte na tlačítko *OK* a PDF dokument uložte.

Licenční politika pro TextSpy

Indexové soubory vytvořené použitím TextSpy Indexeru jsou určeny pouze pro osobní potřebu nebo pro sdílení v rámci firmy. Tyto indexy nejsou určeny k distribuci na veřejně distribuovaných discích CD-ROM nebo na jiných médiích. Chcete-li šířit sadu PDF dokumentů spolu s TextSpy indexem, je třeba získat licenci k šíření indexu a index vytvořit ve verzi TextSpy Indexer CD-ROM Publishing Kit. Více informací získáte na <http://www.textspy.cz>, nebo na <http://www.amsoft.cz> a nebo u svého prodejce.

Znamé problémy

TextSpy nemůže správně vytvořit index a spolehlivě fungovat v dokumentech PDF, ve kterých jsou použita nesprávně lokalizovaná (počeštěná) písma. Tato písma nefungují ani v nových verzích produktů Adobe (InDesign 1.5, Photoshop 5.5 a 6.0, Illustrator 9.0), ani v Microsoft Office 97 a novějších. V tomto případě si prosím vyžádejte od dodavatelů písem opravená písma a s těmi vygenerujte nové soubory PDF. Částečně můžete správnost písem ověřit tak, že v dokumentu PDF textovým nástrojem vyberete text s diakritikou, zkopírujete ho do schránky, vyvoláte okno pro vyhledávání (*Úpravy > Hledat...*) a vložíte text ze schránky. Text by se měl v dialogovém okně zobrazit bez problémů. Nebo uložte PDF soubor do formátu RTF (*Soubor > Uložit jako*), který pak můžete otevřít v příslušné aplikaci.

Vytváření indexu může být náročné na paměť, proto doporučujeme při indexování na Mac OS 9 zvětšit přidělenou paměť pro aplikaci Acrobat.

Technická podpora

V případě potíží se zásuvnými moduly Amos Software TextSpy se prosíme obraťte ne technickou podporu:

Amos Software, spol. s r. o.

Patočkova 61

169 00 Praha 6 – Břevnov

Tel.: +420/2/8401 1222

Fax: +420/2/8401 1212

E-mail: textspy@amsoft.cz

Další informace

Veškeré obchodní a technické informace o produktu Amos Software TextSpy, jeho využití, variantách licencování a cenách, získáte u společnosti Amos Software:

Amos Software, spol. s r. o.

Patočkova 61

169 00 Praha 6 - Břevnov

Tel.: +420/2/84011211

Fax: +420/2/84011212

E-mail: sales@amsoft.cz

<http://www.textspy.cz>

<http://www.textspy.com>

<http://www.amsoft.cz>

© 2000-2001 Amos Software, spol. s r. o. Všechna práva vyhrazena.

Tato příručka, stejně jako v ní popisovaný software, je poskytována na základě licence a může být používána nebo kopírována pouze podle podmínek této licence. Obsah této příručky slouží pouze pro informaci, může se měnit bez upozornění a nelze ho vykládat jako závazek firmy AMOS Software, spol. s r.o. AMOS Software, spol. s r.o. nepřebírá žádnou odpovědnost za chyby nebo nepřesnosti, které se v této příručce mohou objevit. S výjimkami povolenými v licenci nesmí být žádná část této příručky reprodukována, ukládána ve vyhledávacím systému a přenášena v jakékoliv formě nebo jakýmkoliv prostředky, elektronickými, mechanickými záznamovými nebo jinými bez předchozího písemného svolení firmy AMOS Software, spol. s r.o.

TextSpy® je registrovanou ochrannou známkou AMOS Software spol. s r.o. v České republice a jiných zemích.